Method and Apparatus for Multifaceted Profiling of Individual Users
by inventor(s)
Jaideep Srivastava and Priyank Srivastava

5

Field of the Invention

The present invention is in the field of Internet-based services and applications, and pertains more particularly to methods for obtaining a

10 multifaceted profile of a user, which includes a comprehensive set of data obtained from a broad range of resources under a wide array of user categories.

15 ## Background of the Invention

The information network known as the world-wide-web (WWW), which is a subset of the well-known Internet, is arguably the most complete source of publicly accessible information available. Anyone with a suitable

20 Internet appliance such as a personal computer with a standard Internet connection may access (go on-line) and navigate to information pages (termed web pages) stored on Internet-connected servers for the purpose of gathering information and initiating transactions with hosts of such servers and pages.

25 Often times, in order to improve the quality of services offered by a particular website, it is important to understand user activity in relationship to that site. This is to say that while a user is navigating through a website, obtaining a dynamic profile of the user's habits, activities and personal information would prove beneficial to the overall improvement of a service

30 providing or commercial website. In addition to utilizing user profiles for

website service-enhancement purposes, companies routinely pay for such information in order to better target users for advertising and marketing purposes.

In a cobrand relationship known to the inventor, cobrand partners contract with a service-providing entity in order to provide Internet services offered by the entity. The cobranded services are made available to subscribers of the cobrand partners through dedicated servers maintained by the service-providing entity. Users who subscribe to such services typically have at least some personal profile information known to the cobrand partners through their normal subscription and interaction activities. In addition, a service-providing entity may track certain information about users who are accessing and interacting with cobranded services maintained by the service-providing entity. For example, information such as types of products purchased, types of web pages accessed at service sites, frequency of buying, time spent at sites, and so on, may be tracked and stored in a secure database by the service-providing entity. This is made possible by the fact that the service-providing entity maintains and provides the services and the equipment through which the services are made available.

There are a variety of known methods for obtaining information about individual users who visit websites online. Some commonly known methods are sending and retrieving interactive cookies, conducting on-line surveys, parsing completed online forms, recording purchase histories, and many other techniques. A typical user profile automatically compiled by a Web company is limited to information that can be obtained from the user while at one of the company-sponsored sites, or through interacting with the user during registration processes. As such, the profile is not complete or well rounded and tends to reflect content related to the nature of business conducted by the Web company. For example, a purchase history compiled

by a Web-based clothing retailer is limited to the subject of clothing. In order to obtain a well-rounded profile of an individual that covers a variety of topics, information typically must be bought, sold, or traded between Web companies doing business on the Internet. It is known in the art that there are many companies in existence that specialize in information brokering.

In the case of cobranding, where the service-providing entity provides proxy navigation and data summary services for users, data about a user's activity related to interaction with cobrand services includes data related to a plurality of disparate Web-sites, which are involved in some aspect of the cobrand services. It has occurred to the inventor that much information may be automatically obtained about users from user interaction and proxy interaction with many Web sites without being required to obtain the data through purchase or trade with companies hosting Web-sites involved in cobranded services. It is to this aspect of profiling users that the methods and apparatus of the present invention applies.

What is clearly needed is a method that empowers a company to automatically collect data about users either directly or indirectly (through proxy services) whereby the collected data reflects user activity, history, and behavior associated with a plurality of disparate Web-sites. Such a method would enable a company to create dynamic user profiles that are multifaceted and therefore more valuable in the marketplace.

Summary of the Invention

5   According to an embodiment of the present invention, a data-
collection system for collecting data about a user through monitoring user
interaction on a data-packet-network is provided. The system employs a
proxy server connected to the data-packet-network for providing proxy
services and for monitoring user access and interaction with those services.

10   The system also includes at least one dedicated server interface connected to
the data-packet-network for providing user access to the proxy services, and
a software application running on the proxy server for collecting and storing
data obtained as a result of active user-interaction with the proxy services.

        In a preferred embodiment of the present invention, the data-

15   collection system of claim 1 uses the Internet network as a medium of data
transfer. In this aspect, the proxy server and the dedicated server interface
are maintained by a same service-providing entity. The dedicated server
interface is, in preferred application, and Internet file server dedicated to
providing cobrand services to users of a cobrand partner.

20   The data-collection system operates most efficiently in embodiments
wherein the proxy services include a data-collection, aggregation, and
summary service. In preferred aspects the system is operated in a network
environment wherein there are a plurality of dedicated server interfaces,
individual ones of such interfaces dedicated to individual ones of a plurality

25   of participating cobrand partners.

        The software application enabling the system collects demographic
data, account-information data, and on-line behavior data. The collected
data about a user is used to construct a multifaceted user profile. In a

preferred aspect, the data-collection is performed in an entirely automated fashion. However in another aspect additional data obtained through non-automated method is added to the data collected automatically in order to increase the scope of a multifaceted user profile. In preferred aspects, assembly of the multifaceted user profile is automated. Also in preferred aspects, the assembled multifaceted user profile is periodically updated in automated fashion.

According to another aspect of the present invention a method for collecting, storing, and utilizing data related to a user interacting with proxy services on a data-packet-network comprising the steps of (a) monitoring all user activity and transactions associated with interactive use of the proxy services; (b) parsing applicable data resulting from the user activity and transactions; (c) recording the applicable data in a secure an organized fashion; and (d) incorporating the recorded data for the purpose of creating a multifaceted user profile.

In a preferred embodiment, the data-packet-network enabling the method is the Internet network. In this aspect, user transactions described in step (a) include purchases, site registrations, and orders for summary data. In the same aspect, user activity described in step (a) includes activity at an interfacing server. In one aspect of the method, applicable data described in step (c) is recorded to a data repository external to the server recording the data. In another aspect of the method, applicable data is recorded to a data repository held within the server recording the data. In both aspects, the data is used to construct multifaceted user profiles.

Now for the first time, a system and method is provided for automatically collecting data about users interacting on a data-packet-network with provided interactive services in automated fashion wherein the data may be used to construct valuable multifaceted user profiles without

requiring use of manual and other traditional methods for collecting the data and constructing the profiles.

Brief Description of the Drawing Figures

Fig. 1 is an overview of a communication network wherein cobrand user-profiling is practiced according to an embodiment of the present invention.

10   Fig. 2 is a block-diagram illustrating various data categories and data-gathering methods used to create a multifaceted user-profile according to an embodiment of the present invention.

Description of the Preferred Embodiments

15

Fig. 1 is an overview of a communication network 9 wherein multifaceted user profiling is practiced according to an embodiment of the present invention. Communication network 9 contains all of the equipment and network connections required in order to establish a functional network-
20   communication capability.

In a preferred embodiment of the present invention, communication network 9 is implemented on the Internet network, which is represented herein by an Internet backbone 11. Internet backbone 11 represents all of
25   the lines, connection points, and equipment that make up the Internet network as a whole. Therefore, there is no geographic limit to the practice of the present invention.

A plurality of cobrand servers (CBS) 23, 25, and 27 are illustrated, in this example, as connected to Internet backbone 11. CBS 23-27 are adapted as Internet file servers dedicated to users of cobrand services provided by a service-providing company, which also maintains the servers. A main server

5 15, also illustrated as connected to Internet backbone 11, is adapted as a data-aggregation and distribution source and is hosted by the same service provider hosting servers 23-27. Cobrand servers 23-27 are dedicated to cobrand partners and their subscribers whom have elected to access services offered by the service provider, which also hosts server 15.

10 In addition to data-aggregation and distribution services, server 15 also functions, in this example, as an activity-tracking server tracking on-line activity of cobrand users. Companies providing cobrand services to their subscriber-ship typically specialize in providing search capabilities and limited portal services. The companies hosting CBS 23-27 may be termed

15 cobrand partners of the described service-providing company. The nature of services provided by CBS 23-27 will depend on the nature of agreements forged between various cobrand partners and the service-providing entity. For example, some of CBS 23-27 may allow users to obtain e-mail from various e-mail servers. Likewise, each of CBS 23-27 may provide varying

20 functionalities generally related to data collection, aggregation and summary services. It is sufficient to say that subscribers to cobranded services interface with CBS 23-27 in order to receive such services.

A plurality of content servers (CS) 19 and 21 are adapted as Internet data servers hosted by companies contracted to provide specific content to

25 the service-providing entity, which makes the content available through CBS 23-27. Servers 19 and 21 are illustrated herein as connected to Internet backbone 11.

CS 19 and 21 are dedicated to providing specific Web content such as weather information, stock quotes, financial news, entertainment news, and so on. There are many possibilities as to the nature of the content provided by servers 19 and 21. In some cases content provided by servers 19 and 21 may be inaccessible without subscription or membership. It is noted herein that CS 19 and 21 are not, in this example, hosted by the entity hosting main server 15 and cobrand servers 23-27 rather, they are hosted by companies contracting with the service-providing entity of this example and provide content according to contract stipulation.

A plurality of *Free* (FS) content servers 29 and 31 are illustrated herein as connected to Internet backbone 11. FS 29 and 31 in this example are also adapted as Internet data servers, but in this case are not providing content to CBS 23-27. FS 29-31 are not in any way associated with the entity hosting main server 15, however, in some cases may be accessed through main server 15 by proxy such that accessed content may also be tracked by main server 15. Like CS 19-21, there are many possibilities as to the nature of content provided by FS 29-31, the term "free" is used in this example to convey that the services and/or content provided within FS 29-31 is available to anyone who accesses it.

It will be apparent to one with skill in the art that there may be many more CBS, CS, and FS illustrated in this example without departing from the spirit and scope of the present invention. The inventor illustrates only a few of each class of server in this example and deems the illustration sufficient for explanation of the present invention. It is repeated here that CBS 23-27 are cobrand servers maintained by a service-providing entity also hosting main server 15. CS 19-21 are content servers hosted by companies contracting with the described service-providing entity to make their content

available to CBS 23-27. FS 29-31 are free servers not affiliated in any way
with the service-providing entity.

A plurality of cobrand subscribers/users 17 (within dotted rectangle)
is illustrated as having Internet connection to Internet backbone 11. Internet

5   connection in this example includes all of the known methods for accessing
the Internet network. An internet-service-provider (ISP) is not detailed in
this example, but may be assumed to be present in a scenario where users 17
are accessing Internet 11 through normal dial-up modem technology, which
is most common. Other methods include wireless modem connection, cable

10   modem connection, and so on.

In this example, users 17 are illustrated as operating personal
computers (PC) to access Internet 11. In actual practice, any Internet-
capable appliance may be used to practice the present invention as long as it
has network-browsing and display capabilities. Each user 17 may freely

15   navigate to and interact with CBS 23-27, FS 29-30, or CS 19-21. However,
when users are accessing cobrand services from one of CBS 23-27, at least
part of the service enables them to have data obtained from any of CS 19-21
or FS 29-31. For example, if a user 17 is logged into CBS 27 to receive
cobrand services, specific data requested by the user such as weather or

20   financial news would be obtained by proxy from CS 19-21 because of
contract stipulation. If data is requested from any of FS 29-31, then the
requesting user must provide information such as a URL address and type of
data required to enable proxy data collection and presentation because there
is no affiliation between FS 29-31 and CBS 27.

25   Main server 15, as previously described, can log the activities of each
of users 17 when they are interacting with CBS 23-27. Furthermore, server
15 has the ability to record activity information related to any proxy data
request involving any other servers navigated to on behalf of users 17. As a

result, much data specific to a user's activity may be obtained without giving notice to or requesting data from companies hosting servers 19-21 or servers 29-31.

An instance of software (SW) is provided to execute on main server 15. SW at server 15 is adapted to record any user-activity data routed through server 15. Therefore, data obtained through cobrand services offered by CB 23-27 may be logged and identified to particular users, and mined for data to include as profile data. Server 15 may record types of content requested, description and class of items purchased, nature and description of Web-sites targeted for data requests, frequency of same type requests, lists containing URLs and descriptions of user-registered Web sites, and so on. Data about a cobrand user's on-line activity and behavior is compiled and organized within server 15 and then stored as a part of that user's multifaceted profile.

Other information about users may be obtained from companies hosting CBS 23-27. Such information may include personal information related to subscription and registration to receive cobrand services, information obtained through registration and interaction with a cobrand partners regular services routinely accessed by cobrand users, and so on. In this way, much of a multifaceted profile can be automatically generated and stored for cobrand users 17. If a user is a frequent cobrand user and is particularly prolific with on-line activity, then many facets about that user's activity and behavior may be learned and recorded. Profiling a user in this manner greatly reduces the need for trading or purchasing partial profiles compiled by a plurality of un-related sources. However, a multifaceted profile may, of course be enhanced by supplementing the profile with purchased or traded data originating from out-side sources.

Significant market advantages exist for an entity that can obtain a multifaceted profile on each, or at least many of its clients. For example, many user profiles may be mined for more specific data, which may be generated into list-reports detailing preferences and activity traits of many

5 users. Such lists may be created and tailored for advertisement companies or other service providers willing to pay for the information. Thresholds and special rules may also be implemented during profile configuration and maintenance such that profile information may be automatically updated over prolonged user patronage of cobrand services.

10 Fig. 2 is a block-diagram illustrating various data categories and data gathering methods used to create a multifaceted user-profile according to an embodiment of the present invention. In this example, there are three basic categories of data used to create a dynamic multifaceted profile. These categories are illustrated in this example, by element number 37

15 (demographic info), element number 39 (account info), and element number 41 (on-line behavior). Element number 43 represents various data-gathering methods, which may be used to obtain data described in the data categories mentioned above.

Every cobrand user is supplied with a cobrand identification (ID) and

20 a member ID. In this way, any data obtained and warehoused from internal or external sources is easily identifiable to a particular user. Profiles may be automatically assembled using this warehoused data according to enterprise rules. It is important to note herein that each data category 37, 39, and 41 may be populated using automatic data-gathering methods (element 43) such

25 as data capture during interactive sessions. Data capture, as is used in this specification, means recording any and all data about a user during an on-line session, which includes any proxy services. Of course, some data may be supplied by purchasing from the outside, trading, or sharing with partners.

These additional options are included under data-gathering methods 43. For the most part however, automated data capture should be sufficient for supplying a viable multifaceted profile in most cases. This is especially true if a user is prolific in his or her patronage of cobranded services.

5   Referring now to demographic info 37, there is illustrated a plurality of subcategories, which will be discussed from top to bottom. Income level is one important subcategory of demographic information. By knowing a user's income level, advertisements for travel, financial services, and consumables may be more properly tailored for the individual. Location

10   information may also be used to enhance local advertising.

Family relationships are important for understanding lifestyle characteristics, identifying future consumers, and so on. Gender, age, and income levels of family members may also be important in creating a multifaceted profile. Hobbies and other preferences may also be included in

15   data gathered for demographic information. Such data also contributes to understanding lifestyle characteristics and identifying products and services that may fit a user.

Demographic information may be obtained through recording on-line purchase events, registration events, and from general population of on-line

20   forms. Therefore, most demographic information may be obtained through automated data-capturing techniques. In some cases, companies contracting for cobrand service-enhancement with a service-providing entity may simply forward or share some demographic information. Such an arrangement may be, in some embodiments, required as part of contract negotiation. In other

25   cases, especially if certain users are not prolific in on-line activity, demographic information may be purchased from the outside or obtained through trade with outside organizations. Therefore, information that

cannot be obtained through data capture may be obtained through other methods in order to supplement any relatively weak profiles.

Account information 39 is illustrated herein as divided into 4 basic subcategories. These subcategories are listed from top to bottom as entertainment, business, financial, and investment. Each category may be further divided into more subcategories as deemed appropriate. Account information represents data obtained from on-line accounts belonging to a particular user. A user may subscribe to many of these accounts and may add them to cobrand service sites for the purpose of being able to access information from such accounts without physically navigating to them. As proxy services are performed on behalf of a user concerning a user's registered accounts, data used in accessing the accounts and data returned as a result of task performance is collected and incorporated as profile data.

As data is automatically compiled about a user over time, the user's profile becomes more and more valuable and accurate. After a period of time, the service-providing entity maintaining the cobrand services and the user's profile may generate automated reports detailing certain aspects of the user's profile for selective distribution to paying clients. Secure information such as credit card numbers, Social Security numbers, personal identification numbers, passwords, and the like remain in a state of data encryption, or otherwise deleted from data reports containing profile information.

On-line behavior is compiled using user-activity and server-activity data. Such raw data is collected and analyzed in order to compile an on-line behavior profile. Subcategories of user activity that may constitute on-line behavior illustrated within block 41 are listed in discussed from top to bottom.

Identification of Web sites visited either directly or through proxy services maybe automatically captured. Types of products purchased from

those sites may also be automatically captured as well as frequency

parameters associated with purchases illustrated herein as a subcategory of

 products purchased. Frequency of purchases may simply mean how often a

purchase is made over a number of visited sites. The frequency of purchases

5   may also mean the frequency of purchase of one particular product.

Time accessing cobrand services may also be included and

incorporated in forming an on-line behavior profile. Parameters surrounding

banner-ad or sponsor clicking may be incorporated in order to determine

certain preferences. Finally, on-line histories may be created and maintained

10   on virtually any category or subcategory associated with blocks 39 and 41.

In a preferred embodiment of the present invention, most if not all of

the data compiled about a user is collected using automated data capturing

techniques implemented during the normal course of the user accessing

cobrand services. The very nature of such services enables many of these

15   data capturing techniques to be utilized.

It will be apparent to one with skill in the art that there may be more

categories and subcategories included in this example without departing

from the spirit and scope of the present invention. The inventor as outlined

basic categories and basic subcategories and deems them sufficient for

20   illustrative purposes. Therefore, the inclusion of such subcategories and

categories in this example should not be construed as a limitation in any way.

It will also be apparent to one with skill in the art, that the unique

cobrand architecture implemented between the service-providing entity of

Fig. 1 represented by a server 15 and the plurality of cobrand partners

25   utilizing CBS 23-27, which are maintained by the same entity, enables

multifaceted profiling of users to be accomplished in automated fashion.

Therefore, much manual labor and research is eliminated from the data

profiling process.

The method of the present invention may be practiced on any DPN that supports the appropriate Internet protocols. Furthermore, there's no limit to the number of cobrand partners, or end-users that may participate in the practice the present invention. Therefore, the method of the present invention should be afforded the broadest possible scope under examination. The spirit and scope of the present invention is limited only by the claims that follow.

5

10

15

20

25